|  |  |
| --- | --- |
|  | |
| МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ | |
| Федеральное государственное автономное образовательное учреждение высшего образования | |
| **«Дальневосточный федеральный университет» (ДВФУ)** | |
| **Институт математики и компьютерных технологий** | |
| **Департамент информационных и компьютерных систем** | |
| **ОТЧЁТ** | |
| по лабораторной работе №1  «Парная регрессия и корреляция» | |
| по дисциплине «Эконометрика» | |
| Направление «Прикладная информатика в экономике» | |
|  | |
|  | Выполнил студент группы Б9121–09.03.03пиэ  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Н.С. Киптилов |
| Проверил старший преподаватель  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Е. И. Шувалова  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  оценка |
| Г. Владивосток  2024г. | |

# 1. Условие задачи

Используя данные о среднедушевом прожиточном минимуме в месяц одного трудоспособного и среднемесячной заработной платы:

1. Построить поле корреляции.

2. Рассчитать параметры уравнений линейной, гиперболической, степенной, показательной парной регрессии. Записать уравнения в явном виде.

3. Оценить тесноту связи с помощью показателей корреляции и детерминации (для каждого уравнения).

4. Оценить значимость коэффициентов регрессий с помощью t-критерия Стьюдента и доверительных интервалов.

5. Оценить с помощью F-критерия Фишера статистическую надежность результатов регрессионного моделирования.

6. По значениям характеристик, рассчитанных в пунктах 4, 5, выбрать лучшее уравнение регрессии.

7. По лучшему уравнению рассчитать прогнозное значение результата, если прогнозное значение фактора увеличится на 10% от его среднего уровня. Определить доверительный интервал прогноза для уровня значимости =0,05.

# 2. Ход работы

Исходные данные о среднедушевом прожиточном минимуме в месяц одного трудоспособного и среднемесячная заработная плата по субъектам Центрального экономического района на 2022 год были получены на официальном сайте Федеральной службы государственной статистики. Данные представлены на рисунке 1.

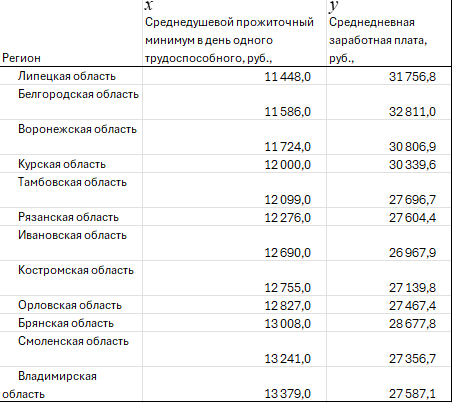


Рисунок 1 – Исходные данные

# 2.1 Поле корреляции

С помощью инструментов MS Excel построено поле корреляции, на котором отображается зависимость среднедушевого прожиточного минимума в месяц одного трудоспособного от среднемесячной заработной платы в субъекте. На поле корреляции так же продемонстрированы линии тренда, соответствующие линейной, логарифмической, степенной и экспоненциальной зависимости, а также уравнения регрессии в явном виде и коэффициент детерминации для каждой зависимости. Поле корреляции представлено на рисунке 2.

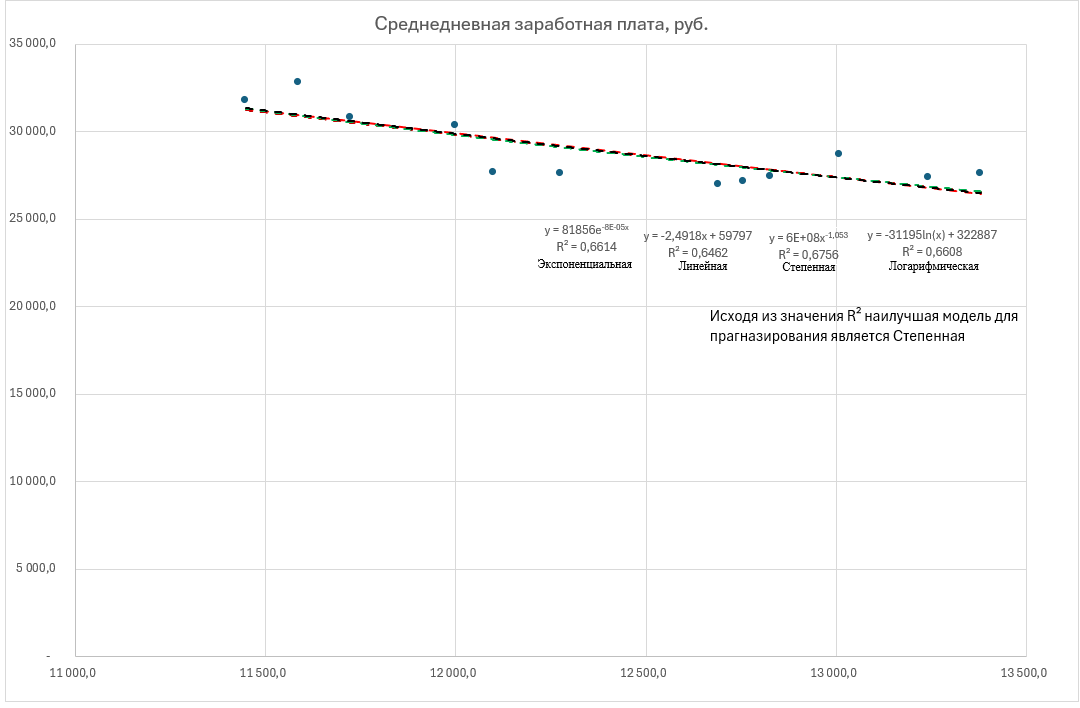


Рисунок 2 – Поле корреляции

# 2.2 Параметры уравнений

Для сравнения парных регрессия были рассчитаны параметры и коэффициенты уравнений, которые представлены на рисунке 3.

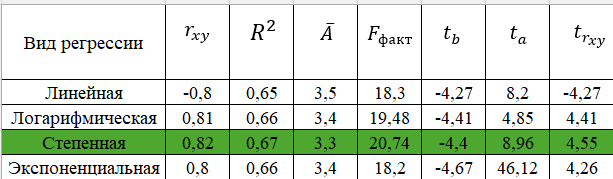


Рисунок 3– Рассчитанные параметры

Из данных расчетов следуют что степенная регрессия наилучшим образом подходит для последующих расчетов.

# 2.3 Оценка тесноты связи

Для оценки параметров уравнения степенной зависимости:

y = a \* xb, которое является нелинейным по фактору x, в уравнении производится замена нелинейной величины. Линеаризация исходного уравнения, представлена на рисунке 4.

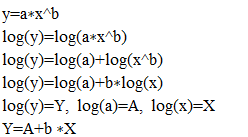


Рисунок 4 – Линеаризация

Для расчета коэффициентов уравнения использовались формулы:

Полученные коэффициенты, а также полученное уравнение в явном виде представлены на рисунке 5.

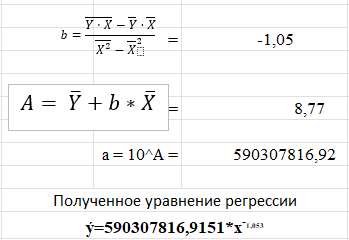


Рисунок 5 – Расчёт параметров a и b

Для оценки тесноты связи были рассчитаны индекс корреляции по формуле:

Для оценки качества построенной модели определен индекс детерминации по формуле: .

А также средняя ошибка аппроксимации по формуле: 

Результаты расчетов представлены на рисунке 6.

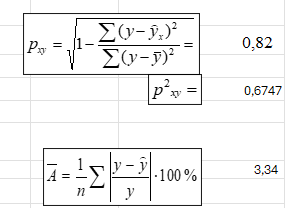


Рисунок 6 – Оценка тесноты связи и качества построенной модели

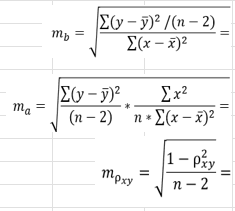
r xy = 0,82 - говорит о сильном характере связи между среднедушевым прожиточным минимумом в месяц одного трудоспособного и среднемесячной заработной платой, т. к. превышает 0,7.

Полученный коэффициент детерминации Rxy²=0,67 говорит о том, что полученное уравнение регрессии хорошо описывает зависимость среднедушевого прожиточного минимума в месяц одного трудоспособного от среднемесячной заработной платы.

Расчёт средней ошибки аппроксимации, которая позволяет получить общее суждение о качестве модели из относительных отклонений по каждому наблюдению. Средняя ошибка аппроксимации Ā = 3.34%, говорит о том, что качество подобранной модели - хорошее, т. к. Ā не превышает 15 %.

# 2.4 Оценка с помощью t-критериев Стьюдента и доверительных интервалов

Для оценки статистической значимости коэффициентов регрессии и корреляции рассчитываются t-критерий Стьюдента и доверительные интервалы каждого из показателей.

Случайные ошибки параметров степенной регрессии и коэффициента корреляции определяются по формулам:

Расчёт случайных ошибок , а также t-критериев приведён на рисунке 7.

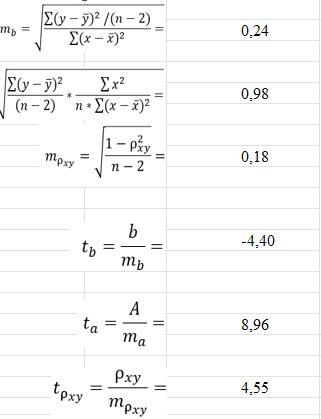


Рисунок 7 – Расчёт случайных ошибок и критериев Стьюдента

Табличное значение t-критериев Стьюдента для числа степеней свободы df = 10 и α = 0,05: tтаб = 2,2281.

ta = 8,96 > t таб = 2,2281

tb = -4,4 < t таб = 2,2281

trxy = 4,55 > t таб = 2,2281

ta > t таб; trxy > t таб, говорит о том что a и trxy сформировались под влиянием систематически действующего фактора x. ta < t таб говорит о случайной природе формирования а.

Для перехода от точечного оценивания к интервальному, с помощью теоремы Гаусса-Маркова было проведено исследование случайных отклонений.

На рисунке 8 представлена зависимость остатков от теоретических значений y. Точки распределены равномерно, что показывает, что остатки являются случайными величинами, и использование метода наименьших квадратов (МНК) оправдано. Теоретические значения ŷx хорошо соответствуют фактическим значениям y. Дисперсия случайной ошибки не меняется в зависимости от номера наблюдения, поэтому условие гомоскедастичности выполняется.

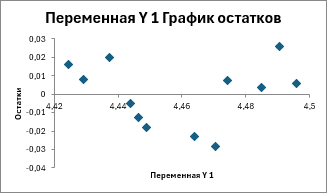


Рисунок 8 – График зависимости остатков от значений y

Построен график зависимости случайных остатков от факторов, включенных в регрессию x. Остатки на графике представлены в виде случайных величин, они не зависят от значений x, что свидетельствует о несмещённости оценок коэффициентов регрессии. Данный график представлен на рисунке 9.



Рисунок 9 - График зависимости остатков от переменной x

График нормального распределения отображает стандартный вид и соответствует ожидаемому распределению. Случайные ошибки на нем распределены нормально. График представлен на рисунке 10.

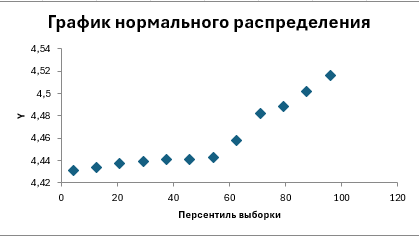


Рисунок 10 – График нормального распределения

Условия Гаусса-Маркова выполняются, оценки параметров регрессии являются несмещёнными, состоятельными и эффективными, что даёт возможность к переходу от точечного оценивания к интервальному.

Для расчета доверительного интервала определяем предельную ошибку Δ для каждого показателя:, .

Результат представлен на рисунке 11.



Рисунок 11 - Расчёт предельных ошибок показателей a и b

Доверительные интервалы показателя a, рассчитанные с помощь предельной ошибки, представлены на рисунке 12.

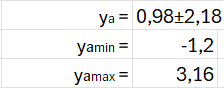


Рисунок 12 – Расчёт доверительных интервалов показателя a

Доверительные интервалы показателя b, рассчитанные с помощь предельной ошибки, представлены на рисунке 13.

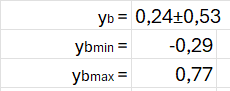


Рисунок 13 – Расчёт доверительных интервалов показателя b

Так как в границы доверительного интервала не попадает ноль, т. е. нижняя и верхняя граница отрицательны для параметра а или нижняя и верхняя границы положительны для параметра b, то оцениваемые параметры a и b не принимают нулевых значений, т. е. являются статистически значимыми и существенно отличны от нуля.

# 2.5 Оценка с помощью F-критерия Фишера

Расчет критерия Фишера (F-критерия), который является показателем надежности и значимости уравнения в целом, по формуле: 

представлен на рисунке 15.



Рисунок 15 – Критерий Фишера

19,48 так как , уравнение надежно, значимо, показатель тесноты связи надежен и отражает устойчивую зависимость среднедушевого прожиточного минимума в месяц на одного трудоспособного от среднемесячной заработной платы.

# 2.6 Прогнозирование

yp определяется путем подстановки в уравнение регрессии = a \* xb соответствующего (прогнозного) значения xp. Полученные yp и xp представлены на рисунке 15.



Рисунок 15 - Расчёт прогнозных значений

Результат расчёта остаточной дисперсия, используя формулу: , а также расчёта средней стандартной ошибки прогноза , используя формулу: представлен на рисунке 17.

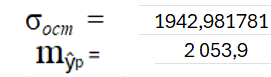


Рисунок 17 - Расчёт остаточной дисперсии и средней стандартной ошибки

Предельная ошибка прогноза, которая в 95% случаев не будет превышена, представлена на рисунке 18.



Рисунок 18 – Расчёт предельной ошибки прогноза

Расчёт доверительного интервала прогноза представлен на рисунке 18.

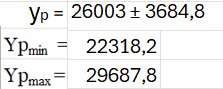


Рисунок 18 – Расчёт доверительного интервала прогноза

# 3. Полученные результаты

В результате выполнения данной лабораторной работы при использовании данных о среднедушевом прожиточном минимуме в месяц одного трудоспособного и среднемесячной заработной платы были выполнены следующие задачи:

1. Построено поле корреляции.
2. Рассчитаны параметры уравнений линейной, гиперболической, степенной, показательной парной регрессии. Записаны уравнения в явном виде.
3. Произведена оценка тесноты связи с помощью показателей корреляции и детерминации (для каждого уравнения).
4. Произведена оценка значимости коэффициентов регрессий с помощью t-критерия Стьюдента и доверительных интервалов.
5. проведено исследование случайных отклонений с помощью метода наименьших квадратов и теоремы Гаусса-Маркова;
6. Произведена оценка с помощью F-критерия Фишера статистической надежности результатов регрессионного моделирования.
7. Выбрано лучшее уравнение регрессии.
8. По лучшему уравнению рассчитано прогнозное значение результата, если прогнозное значение фактора увеличится на 10% от его среднего уровня. Определен доверительный интервал прогноза для уровня значимости =0,05.