Лабораторная работа №1

Кориба М. Н.

13 11 2020

# Раздел 1.

В этом разделе необходимо собрать данные по социально-экономическим показателям в регионах РФ и изучить их взаимосвязи.

# **Необходимо построить:**

* Фактические плотности распределения переменных с наложенными теоретическими плотностями нормального закона (разместить графики на одном полотне)..
* Тест Андерсона-Дарлинга на нормальность: таблица со статистикой теста, P-значением и выводом по каждой переменной.
* Точечные графики взаимного разброса (?pairs), сделать маркеры полупрозрачными.
* Взаимосвязь показателей: графическое представление корреляционной матрицы с помощью функции corrplot(). Незначимые коэффициенты корреляции скрыть, остальные показать квадратами.

## Обозначение показателей:

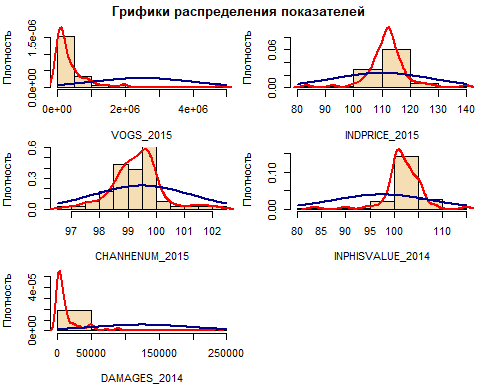
* *VOGS\_2015* - Объем отгруженных товаров собственного производства, выполненных работ и услуг собственными силами по видам экономической деятельности: обрабатывающие производства
* *INDPRICE\_2015* - Индексы цен производителей промышленных товаров по видам экономической деятельности: обрабатывающие производства
* *CHANHENUM\_2015* - Изменение среднегодовой численности занятых
* *INPHISVALUE\_2014* -Индекс физического объема ВРП
* *DAMAGES\_2014* - Сумма убытка организаций по отдельным видам экономической деятельности: обрабатывающие производства

## Описательная статистика

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | VOGS\_2015 | INDPRICE\_2015 | CHANHENUM\_2015 | INPHISVALUE\_2014 | DAMAGES\_2014 |
| Среднее | 402737.3 | 111.9 | 99.3 | 101.8 | 16230.2 |
| Стандартное отклонение | 643174.9 | 7.0 | 0.9 | 3.8 | 31119.3 |
| Коэффициент вариации, % | 159.7 | 6.3 | 0.9 | 3.7 | 191.7 |

## Анализ распределния данных

Построим график



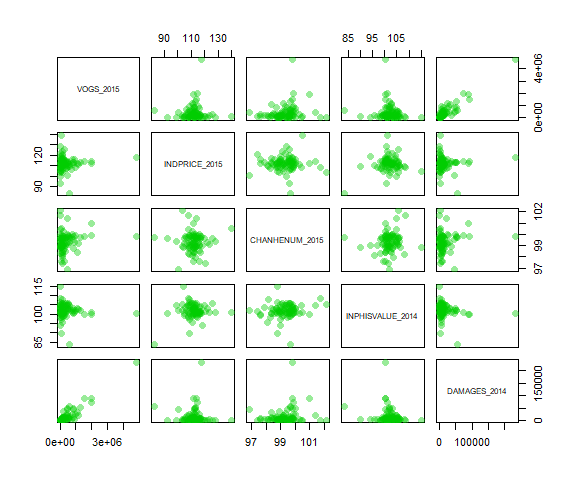
## Тест Андерсона-Дарлинга на нормальность

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | VOGS\_2015.A | INDPRICE\_2015.A | CHANHENUM\_2015.A | INPHISVALUE\_2014.A | DAMAGES\_2014.A |
| Статистика теста | 9.0173 | 2.4155 | 1.2759 | 2.5441 | 11.78 |
| p-значение | 0.0000 | 0.0000 | 0.0024 | 0.0000 | 0.00 |

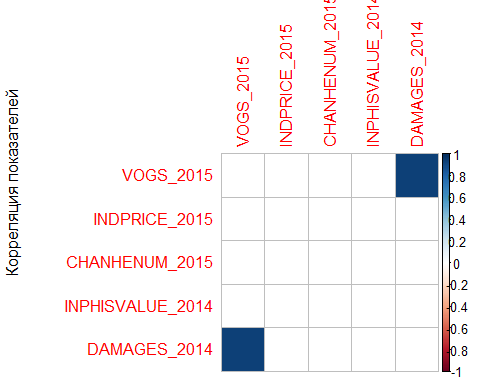
**Вывод:** Проверяя гипотезу нормального распределения для переменных на уровне значимости 0.05, мы отвергаем все показатели.

## Анализ взаимосвязей показателей

### Графики разброса



### Корреляционная матрица



**Вывод:** Единственная значимая корреляция наблюдается между *‘VOGS\_2015’* и *‘DAMAGES\_2014’*.

# Раздел 2.

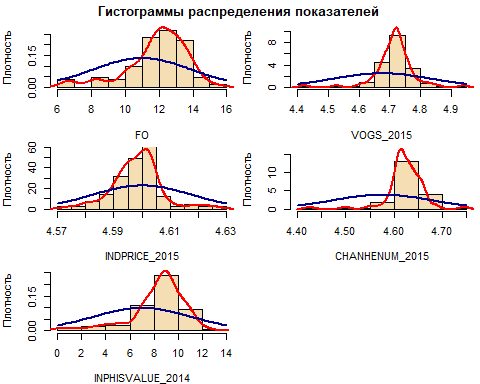
В этом разделе необходимо произвести те же самые действия, с логарифмированными исходными данными.

## Описательная статистика логарифмированных данных:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | VOGS\_2015 | INDPRICE\_2015 | CHANHENUM\_2015 | INPHISVALUE\_2014 | DAMAGES\_2014 |
| Среднее | 11.9 | 4.7 | 4.6 | 4.6 | 8.3 |
| Стандартное отклонение | 1.8 | 0.1 | 0.0 | 0.0 | 2.2 |
| Коэффициент вариации, % | 15.1 | 2.1 | 0.0 | 0.0 | 26.5 |

## Анализ распределния логарифмированных данных

Строим фактическую плотность распределения переменных с наложенными теоретическими плотностями нормального закона



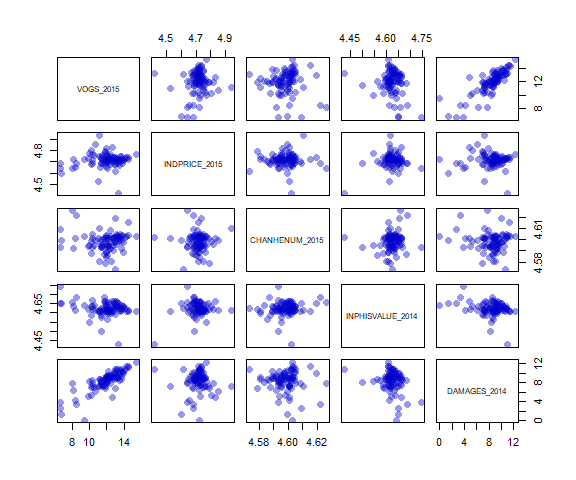
## Тест Андерсона-Дарлинга на нормальность для логарифмированных данных

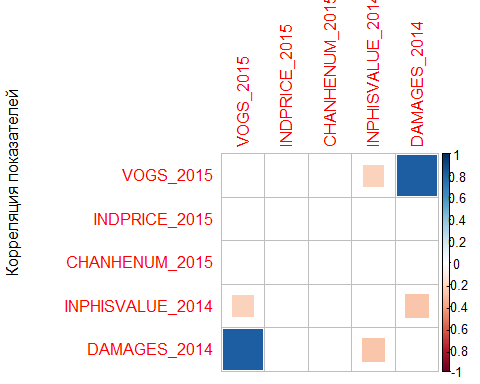
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | VOGS\_2015.A | INDPRICE\_2015.A | CHANHENUM\_2015.A | INPHISVALUE\_2014.A | DAMAGES\_2014.A |
| Статистика теста | 1.9148 | 2.6327 | 1.2690 | 2.9098 | 2.3947 |
| p - значение | 0.0001 | 0.0000 | 0.0025 | 0.0000 | 0.0000 |

**Вывод:** По итогам теста Андерсона-Дарлинга на нормальность распредления логарифмированных переменных, при уровне значимости 0,05, отвергаются все переменные.

## Анализ взаимосвязей логарифмированных показателей

### Графики разброса логарифмированных величин





**Вывод:** Наблюдается сильная корреляционная связь между показателями *‘VOGS\_2015’* и *‘DAMAGES\_2014’*, и значимая корреляция наблюдается между показателями *‘VOGS\_2015’* и *‘INDPHISVALUE\_2014’*, *‘INDPHISVALUE\_2014’* и *‘DAMAGES\_2014’*.