Лабораторная работа № 1

Ямпольский Антон

21 12 2020

## Раздел I.

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

**Необходимо построить:**

* гистограммы на одном полотне, отразив теоретические плотности распределения;
* таблицу со статистикой теста Андерсона-Дарлинга на нормальность;
* точечные графики взаимного разброса с полупрозрачными маркерами;
* графическое представление корреляционной матрицы.

### Обозначения количественных показателей:

* *IPI.2013* – Индексы промышленного производства.
* *PIM.2013* – Индесы цен производителей промышленных товаров по видам экономической деятельности: обрабатывающие производства.
* *DDFA.2013* – Степень износа основных фондов.
* *FCI.2012* – Инвестиции в основвной капитал на душу населения.
* *DLR.2013* – Задолжность по кредитам в рублях, предоставленым кредитными организациями юридическим лицам. ## Импорт данных:

#Файл с данными содержит 92 строк и 8 столбцов.

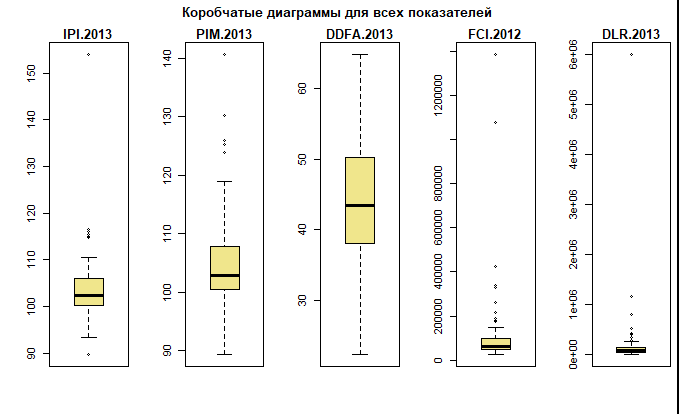
## Описательная статистика

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IPI.2013 | PIM.2013 | DDFA.2013 | FCI.2012 | DLR.2013 |
| Среднее | 103.6 | 105.1 | 43.9 | 114737.6 | 194488.6 |
| Стандартное отклонение | 7.5 | 8.3 | 9.0 | 191500.2 | 666842.3 |
| Коэффициент вариации, % | 7.2 | 7.9 | 20.5 | 166.9 | 342.9 |

**Вывод:** показатели *FCI.2012*, *DLR.2013* неоднородны, т.к. коэффициент вариации > 33%

## Анализ распределения данных

# Коробчатые диаграммы



**Вывод:** Можно заметить, что каждый показатель имеет выбросы (свои аномальные точки)

# Тесты на нормальность

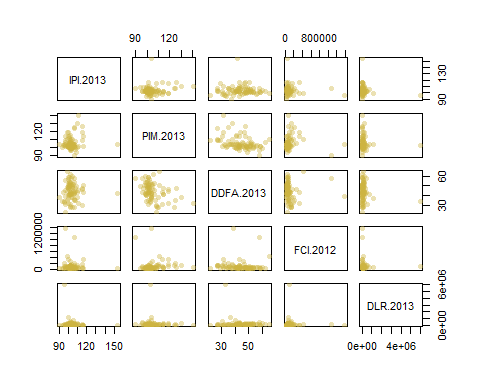
# Тест Андерсона-Дарлинга

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IPI.2013.A | PIM.2013.A | DDFA.2013.A | FCI.2012.A | DLR.2013.A |
| Статистика теста | 4.2887 | 4.797 | 0.2054 | 16.6467 | 21.1813 |
| Критическое значение | 0.0000 | 0.000 | 0.8676 | 0.0000 | 0.0000 |

**Вывод:**Проверяя гипотезу нормального распределения для переменных на уровне значимости 0.05, мы отвергаем все факторы,так как их p-значения < 0.05.

# Анализ линейных взаимосвязей

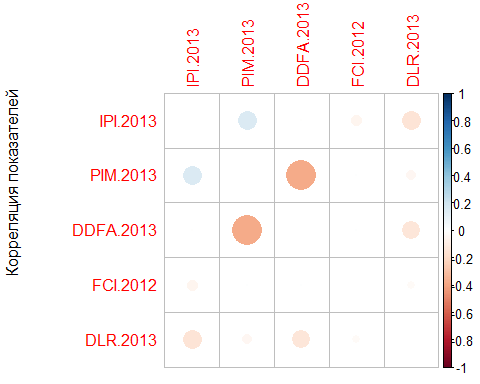
# Точечные графики разброса



**Вывод:** по первой строке графиков попытаемся определить наличие связи между *IPI.2013* и остальными показателями. У *IPI.2013* и *PIM.2013* и *DDFA.2013* положительная связь. А с остальными у *IPI.2013* связи нет.

# Корреляционная матрица

# коэффициенты Пирсона с Р-значениями

 **Вывод:** Взаимосвязь *IPI.2013* слабо значима с показателем *PIM.2013*, а с *DDFA.2013* наблюдается обратная связь.

## Раздел II

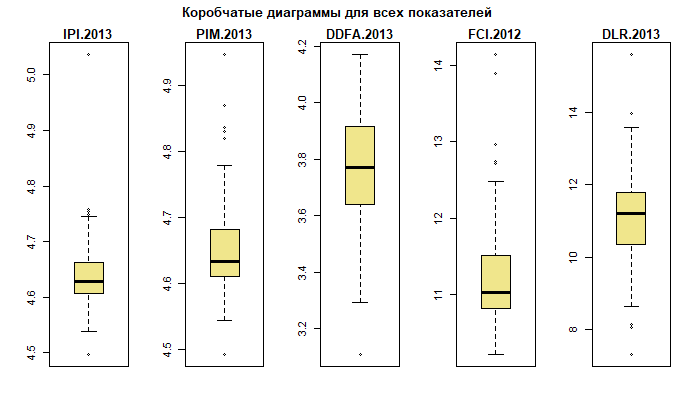
В этом разделе необходимо прологарифмировать данные из первого раздела и провести с новыми данными те же махинации. ## Описательная статистика логарифмированных данных:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IPI.2013 | PIM.2013 | DDFA.2013 | FCI.2012 | DLR.2013 |
| Среднее | 4.638 | 4.652 | 3.761 | 11.243 | 11.122 |
| Стандартное отклонение | 0.066 | 0.075 | 0.217 | 0.725 | 1.329 |
| Коэффициент вариации, % | 1.423 | 1.612 | 5.770 | 6.448 | 11.949 |

**Вывод:** показатели считаются не полностью однородными, т.к. ни один коэффициент вариации не превышает 33%.

## Анализ распределения данных

#Построим коробчатые диаграммы.



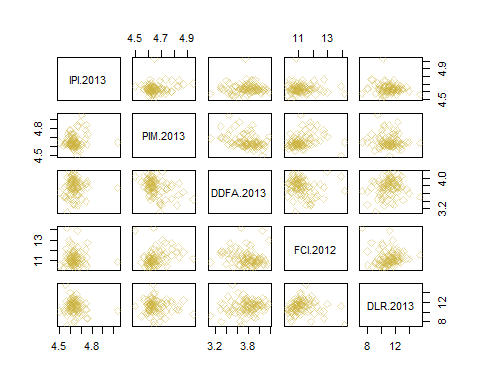
**Вывод:** Каждый признак имеет выбросы.

## Тест Андерсона-Дарлинга на нормальность распределения

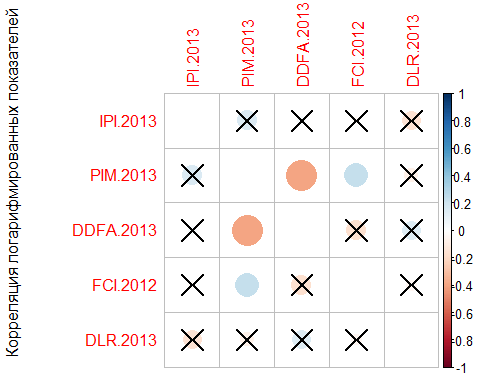
|  |  |  |
| --- | --- | --- |
|  | p | W |
| IPI.2013 | 0.0000 | 2.9886 |
| PIM.2013 | 0.0000 | 3.9995 |
| DDFA.2013 | 0.0377 | 0.7956 |
| FCI.2012 | 0.0000 | 2.7965 |
| DLR.2013 | 0.1633 | 0.5379 |

**Вывод:** Проверяя гипотезу нормального распределения для переменных на уровне значимости 0.05, мы отвергаем все, кроме *“DLR.2013”*, так как остальные p-значения < 0.05 .

## Графики разброса



**Вывод:** по первой строке графиков попытаемся определить наличие связи между *IPI.2013* и остальными показателями. У *IPI.2013* и *FCI.2012* и *DLR.2013* положительная связь. А с остальными у *IPI.2013* связи нет. ## Корреляционная матрица



**Вывод:** Взаимосвязь *DDFA.2013* имеет обратную свзь с показателем *PIM.2013*, а *FCI.2012* слабо значима с показателем *DDFA.2013*.