Лабораторная работа №1

Кравчук Д. А.

27 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:

* *IndProd* – Индексы промышленного производства.
* *PromPrices* – Индесы цен производителей промышленных товаров по видам экономической деятельности: обрабатывающие производства.
* *FundsDeter* – Степень износа основных фондов.
* *CapInvests* – Инвестиции в основвной капитал на душу населения.
* *CreditsIndebted* – Задолжность по кредитам в рублях, предоставленым кредитными организациями юридическим лицам. ## Импорт данных:

#Файл с данными содержит 92 строк и 8 столбцов.

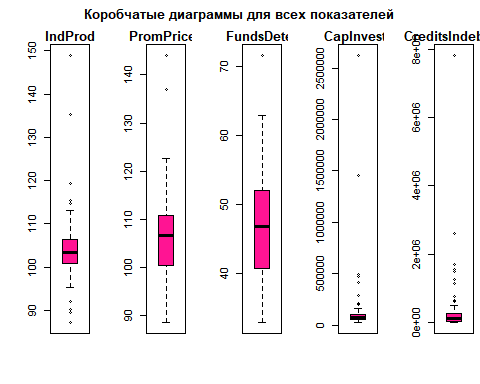
## Описательная статистика

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IndProd | PromPrices | FundsDeter | CapInvests | CreditsIndebted |
| Среднее | 104.2 | 106.0 | 47.2 | 143752.7 | 352510.4 |
| Стандартное отклонение | 8.1 | 9.1 | 8.0 | 327184.1 | 937483.2 |
| Коэффициент вариации, % | 7.8 | 8.6 | 16.9 | 227.6 | 265.9 |

**Вывод:** показатели *CapInvests*, *CreditsIndebted* неоднородны, т.к. коэффициент вариации > 33%

## Анализ распределения данных

# Коробчатые диаграммы



**Вывод:** Можно заметить, что каждый показатель имеет выбросы

# Тесты на нормальность

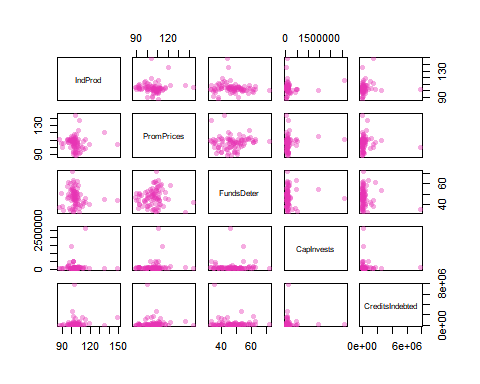
# Тест Андерсона-Дарлинга

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IndProd.A | PromPrices.A | FundsDeter.A | CapInvests.A | CreditsIndebted.A |
| Статистика теста | 4.8583 | 1.4760 | 0.2713 | 20.0582 | 17.162 |
| Критическое значение | 0.0000 | 0.0008 | 0.6649 | 0.0000 | 0.000 |

**Вывод:**Проверяя гипотезу нормального распределения для переменных на уровне значимости 0.05, мы отвергаем все факторы, кроме *FundsDeter* ,так как его p-значение = 0,7875, а остаьные p-значения < 0.05.

# Анализ линейных взаимосвязей

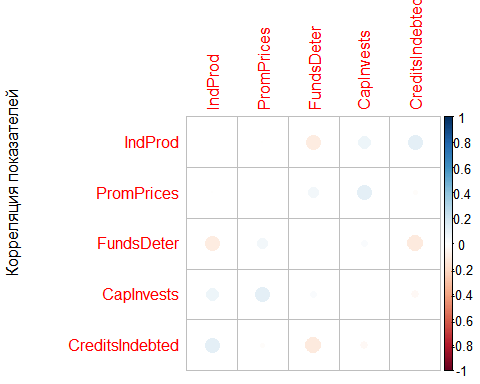
# Точечные графики разброса



**Вывод:** по первой строке графиков попытаемся определить наличие связи между *IndProd* и остальными показателями. У *IndProd* и *PromPrices* положительная связь. А с остальными у *IndProd* связи нет.

# Корреляционная матрица

# коэффициенты Пирсона с Р-значениями

 **Вывод:** Взаимосвязь *IndProd* слабо значима с показателем *CapInvests*, а с *CreditsIndebted* наблюдается обратная связь.

## Раздел II

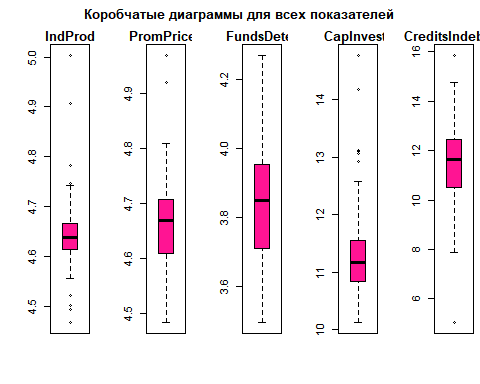
В этом разделе необходимо прологарифмировать данные из первого раздела и провести с новыми данными те же махинации. ## Описательная статистика логарифмированных данных:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IndProd | PromPrices | FundsDeter | CapInvests | CreditsIndebted |
| Среднее | 4.643 | 4.660 | 3.840 | 11.310 | 11.573 |
| Стандартное отклонение | 0.072 | 0.084 | 0.170 | 0.801 | 1.592 |
| Коэффициент вариации, % | 1.551 | 1.803 | 4.427 | 7.082 | 13.756 |

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

## Анализ распределения данных

#Построим коробчатые диаграммы.



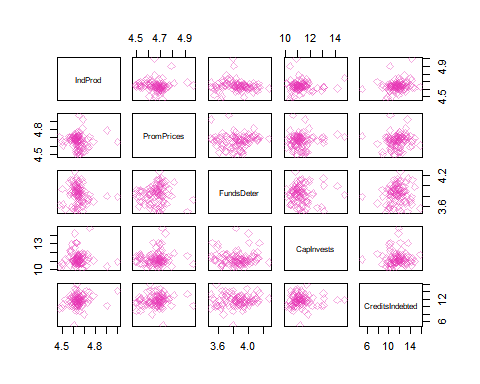
**Вывод:** Каждый признак имеет выбросы.

## Тест Андерсона-Дарлинга на нормальность распределения

|  |  |  |
| --- | --- | --- |
|  | p | W |
| IndProd | 0.0000 | 3.8266 |
| PromPrices | 0.0014 | 1.3706 |
| FundsDeter | 0.5147 | 0.3268 |
| CapInvests | 0.0000 | 3.5359 |
| CreditsIndebted | 0.1169 | 0.5944 |

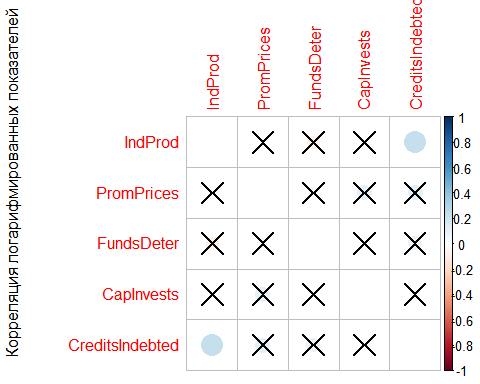
**Вывод:** Проверяя гипотезу нормального распределения для переменных на уровне значимости 0.05, мы отвергаем все, кроме *“FundsDeter”* и *“CreditsIndebted”*, так как остальные p-значения < 0.05 .

## Графики разброса



**Вывод:** по первой строке графиков попытаемся определить наличие связи между *IndProd* и остальными показателями. У *IndProd* и *PromPrices* положительная связь. А с остальными у *IndProd* связи нет.

## Корреляционная матрица



**Вывод:** Взаимосвязь *FundsDeter* значима с показателем *PromPrices*, и при этом она стала сильнее.