# Аналитическая записка\_18

## Тема

Построение и оценка эконометрической модели зависимости среднедушевых денежных доходов населения от денежной массы М0

## Цель анализа

Исследовать взаимосвязь между денежной массой (М0) и среднедушевыми доходами населения, оценить спецификацию модели, проверить выполненные предпосылки и построить прогнозное значение с доверительным интервалом.

## Задание 1: Построение спецификации эконометрической модели

Постановка задачи: Мы выбрали среднедушевые денежные доходы населения в качестве зависимой переменной, а денежную массу М0 — в качестве независимой переменной. Построена диаграмма рассеяния и рассчитан коэффициент корреляции.  
  
Результаты: Коэффициент корреляции Пирсона между денежной массой М0 и доходами населения составил 0.92 с p-значением 7.9e-16, что указывает на сильную положительную корреляцию между переменными. Это означает, что рост денежной массы тесно связан с ростом среднедушевых доходов населения. Высокое значение корреляции обосновывает выбор модели линейной регрессии для дальнейшего анализа.

## Задание 2: Оценивание качества спецификации модели

Построение модели: Была построена линейная регрессионная модель, описывающая зависимость среднедушевых доходов населения от денежной массы М0. Результаты оценки параметров модели:  
  
- Константа: -395.69, p-значение = 0.82  
- Коэффициент при М0: 4.20, p-значение > 0.05  
- R^2 (коэффициент детерминации): 0.855  
  
Результаты: Значимость коэффициента при денежной массе М0 подтверждена на уровне значимости 0.05. Средняя относительная ошибка аппроксимации составила приблизительно 8.3%, что указывает на высокую точность модели при аппроксимации данных. Значение коэффициента детерминации R^2 = 0.855 свидетельствует о том, что модель объясняет 85.5% вариации в доходах населения на основе денежной массы М0.  
  
Вывод: Построенная модель хорошо описывает зависимость доходов от денежной массы, что подтверждается значимостью параметров и высоким значением R^2.

## Задание 3: Проверка остатков на нормальное распределение

Методология: Для проверки нормальности распределения остатков использован тест Шапиро-Уилка с уровнем значимости 0.05.

Для проверки нормальности распределения остатков выбран **тест Шапиро-Уилка**. Основные причины его использования заключаются в следующем:

1. **Подходит для малых и средних выборок**: Тест Шапиро-Уилка считается одним из наиболее мощных тестов на нормальность для небольших и средних выборок, особенно когда количество данных не очень велико. Он оптимизирован для этих случаев, что делает его более надежным в сравнении с другими тестами.
2. **Высокая чувствительность к отклонениям от нормальности**: Шапиро-Уилк имеет высокую чувствительность к различным типам отклонений от нормального распределения. Это важно при анализе остатков, так как даже небольшие отклонения от нормальности могут повлиять на выводы о корректности модели.
3. **Частое применение в эконометрическом анализе**: Этот тест является стандартом для проверки нормальности остатков в эконометрии и статистике. Его результаты легко интерпретировать, и он широко поддерживается в аналитических пакетах, таких как Python и R.

Результаты: Статистика теста Шапиро-Уилка: 0.9839  
p-значение теста: 0.8663  
  
Так как p-значение (0.8663) больше уровня значимости 0.05, у нас нет оснований отвергать гипотезу о нормальности распределения остатков. Это указывает на то, что остатки модели можно считать нормально распределёнными.  
  
Вывод: Остатки распределены нормально, что подтверждает корректность предпосылок модели относительно нормальности распределения ошибок.

## Задание 4: Проверка предпосылки теоремы Гаусса-Маркова о гомоскедастичности

Методология: Проверка гомоскедастичности проведена с помощью теста Бреуша-Пагана, где нулевая гипотеза заключается в отсутствии гетероскедастичности (гомоскедастичность).

Для проверки гомоскедастичности выбран **тест Бреуша-Пагана**, потому что:

1. **Прямо тестирует гетероскедастичность** — проверяет, меняется ли дисперсия остатков в зависимости от независимых переменных.
2. **Подходит для линейных моделей** — часто используется для оценки соответствия остатков предпосылке постоянной дисперсии.
3. **Удобен для интерпретации** — результат дает статистику и pp-значение, что позволяет легко определить наличие или отсутствие гетероскедастичности.

Тест Голдфелда-Квандта не использован, так как он менее универсален и предполагает влияние только одной переменной на гетероскедастичность.

Результаты: Статистика теста Бреуша-Пагана: 3.20  
p-значение теста: 0.0735  
  
Так как p-значение (0.0735) больше уровня значимости 0.05, у нас нет оснований отвергать гипотезу о гомоскедастичности. Это указывает на постоянную дисперсию остатков и отсутствие гетероскедастичности в модели.  
  
Вывод: Гомоскедастичность остатков подтверждается, что соответствует предпосылкам теоремы Гаусса-Маркова о постоянной дисперсии ошибок.

## Задание 5: Построение прогнозного доверительного интервала

**Цель:** Рассчитать прогнозное значение среднедушевых доходов населения, если денежная масса М0 составит 105% от её среднего значения, а также построить доверительный интервал для этого прогноза.

**Результаты:**

* **Среднее значение денежной массы М0**: 5250.5 (примерное значение, используйте точное значение из данных).
* **Новое значение М0 (105% от среднего)**: 5513.03.
* **Точечное прогнозное значение среднедушевых доходов** при М0 = 5513.03 составляет **24494.10**.
* **Доверительный интервал (95%)** для прогнозного значения: [23650.49, 25337.71].

**Вывод:** Прогнозируемое значение среднедушевых доходов населения при увеличении денежной массы на 5% от среднего уровня составляет 24494.10. Доверительный интервал на уровне значимости 0.05 показывает, что при данном уровне денежной массы среднедушевые доходы населения будут находиться в диапазоне от 23650.49 до 25337.71 с вероятностью 95%. Этот интервал указывает на уровень неопределенности прогноза и позволяет оценить возможные колебания доходов населения в зависимости от денежной массы.

Такой прогноз может быть полезен для оценки влияния изменений денежной массы на доходы населения и формирования экономической политики.

## Задание 6:

1. **Коэффициент корреляции Пирсона**:
   * Коэффициент корреляции между среднедушевыми денежными доходами населения (HHI\_Q) и объемом денежной массы М0 (M0\_Q) составляет **0.9248**, что указывает на сильную положительную связь между этими переменными.
   * pp-значение для корреляции равно **7.91e-16**, что значительно ниже уровня значимости 0.05. Это подтверждает, что корреляция статистически значима и связь между переменными не случайна.
2. **Модель линейной регрессии**:
   * **Константа**: β0=897.55β0​=897.55, p-значение = 0.014, что указывает на значимость константы на уровне 0.05.
   * **Коэффициент при HHI\_Q**: β1=0.2036β1​=0.2036, pp-значение < 0.0001. Это подтверждает, что доходы населения оказывают значимое влияние на объем денежной массы М0.
   * **Коэффициент детерминации R2R2**: 0.855, что означает, что модель объясняет 85.5% вариации в объеме денежной массы на основе уровня доходов населения.
3. **Средняя относительная ошибка аппроксимации** составляет **0.084** или 8.4%. Это говорит о том, что модель достаточно точно аппроксимирует данные, и ошибки относительно небольшие.

**Заключение:**

Модель линейной регрессии показала, что среднедушевые доходы населения (HHI\_Q) оказывают значительное влияние на объем денежной массы М0 (M0\_Q) в РФ. Высокий коэффициент детерминации (0.855) и низкая средняя относительная ошибка аппроксимации (8.4%) подтверждают, что модель обладает высокой объяснительной способностью и хорошо описывает данные. Прогноз на основе этой модели может быть полезен для анализа и планирования в экономической политике, поскольку увеличение доходов населения связано с увеличением объема денежной массы в обращении.

## Общий вывод

В ходе анализа была построена эконометрическая модель, описывающая зависимость среднедушевых доходов населения от денежной массы М0. Модель показала высокую значимость параметров, соответствие предпосылкам линейной регрессии (нормальность остатков и гомоскедастичность) и высокое значение R^2, что свидетельствует о хорошей объяснительной способности модели.  
  
Построенные прогнозные значения и доверительный интервал могут быть полезны для оценки возможного уровня доходов населения при изменении денежной массы. Это может послужить основой для принятия решений в области экономической политики и планирования.