

ЛЕКЦИЯ 4. F-тест: как понять, значима ли ваша модель вообще

Введение

На предыдущих занятиях вы научились строить регрессионные модели.

Теперь вы знаете, как:

формулировать гипотезу,

собирать данные,

преобразовывать ответы в числа,

строить линейную и полиномиальную регрессию,

интерпретировать коэффициенты.

Но наступает критический момент в развитии аналитического мышления:

“А насколько вообще **имеет смысл** моя модель?”

“А может быть, все эти коэффициенты — просто шум и случайность?”

Если вы не задаёте себе этот вопрос — вы **не проверяете гипотезу**, а просто **подставляете числа**.

 **Почему этого недостаточно?**

Допустим, вы построили модель, получили:

$$y = 50 + 5.2x - 0.3x^2$$

Отлично. Но:

Насколько это уравнение **лучше**, чем просто угадать среднее значение y ?

А если бы вы случайным образом раскидали данные — вы бы тоже получили такое уравнение?

Какова **общая значимость** всех переменных?

Вот здесь вступает в игру F-тест.

Это статистическая проверка того, что **вся ваша модель в целом имеет смысл**, а не только отдельные её коэффициенты.

Цель F-теста

F-тест отвечает на вопрос:

"Объясняет ли хотя бы **одна из переменных** значимую часть изменений в y ?"

Если ответ **да** — модель значима.

Если **нет** — она может не быть лучше случайного угадывания.

Формулировка гипотезы для F-теста

H_0 (нулевая гипотеза):

Модель не объясняет y . Все коэффициенты = 0 (кроме b_0).

H_1 (альтернатива):

Хотя бы один коэффициент $\neq 0 \rightarrow$ модель **объясняет** y .



Формула F-статистики

Технически, она считается так:

$$F = (SSR / k) / (SSE / (n - k - 1))$$

Где:

SSR — сумма квадратов регрессии (объяснённая часть),

SSE — сумма квадратов остатков (необъяснённая часть),

k — количество объясняющих переменных,

n — количество наблюдений.

Но... вам не нужно считать это вручную!

F-тест рассчитывается автоматически, когда вы используете LINEST в Excel или Google Sheets с параметрами TRUE, TRUE.

Как сделать F-тест в Google Sheets или Excel

Допустим, у вас таблица:

Респондент	Часы подготовки (x)	x^2	Оценка (y)
1	1	1	55
2	3	9	70
3	5	25	85
4	7	49	88
5	10	100	82
6	12	144	75
7	15	225	60

Вы добавили x^2 для полиномиальной модели.

Формула в Google Sheets:

```
=LINEST(D2:D8, A2:B8, TRUE, TRUE)
```

Формула в Excel (русский язык):

```
=ЛИНЕЙН(D2:D8; A2:B8; ИСТИНА; ИСТИНА)
```

Возвращается **таблица**, где одна из ячеек содержит **значение F**.

Также будет возвращено значение R^2 , стандартные ошибки, значения коэффициентов.

Интерпретация F-значения

F-значение	Что это значит
Высокое (например, > 4–5)	Модель статистически значима
Близкое к 1	Модель может не иметь смысла
< 1	Модель хуже, чем просто угадать среднее значение y

Иногда Google Sheets или Python дают вам не только F, но и p-value.

p-value	Интерпретация
< 0.05	Модель значима (95% уверенность)
> 0.05	Недостаточно оснований доверять модели

Пример: интерпретация результата

Допустим, *LINEST* вернул:

$F = 12.5$

$R^2 = 0.91$

Вывод:

Модель объясняет **91%** изменений в y

Вероятность, что такая точность — случайна, очень мала

Модель значима, можно переходить к анализу отдельных коэффициентов

Частые ошибки

Смотреть только на R^2 , игнорируя F

Интерпретировать коэффициенты без уверенности в значимости модели

Не использовать TRUE, TRUE в формуле *LINEST*, теряя F -статистику

Применять F -тест к простой регрессии без нужды (там проще t -тест)

Использование ИИ

Разрешено:

Инструмент	Что можно делать
ChatGPT / Data Analyst	Объяснит, что такое F , как интерпретировать
Excel Copilot	Автоматически посчитает F и p -value
Google Sheets Explore	Покажет статистику регрессии

Категорически запрещено:

Не проверять модель перед интерпретацией

Копировать “хорошие” значения F из интернета

Писать “модель значима”, не видя значения F

Использовать F только “для галочки”

Вывод

F -тест — это фильтр.

Он говорит вам: “Эта модель заслуживает вашего внимания” или “Нет, она не лучше случайного предположения”.

Если F-тест показывает значимость — вы можете спокойно перейти к анализу коэффициентов и строить выводы.

Если нет — лучше пересмотреть модель или вернуться к гипотезе.